University of Rochester

William E. Simon Graduate School of
Business Administration

The Bradley Policy Research Center
Financial Research and Policy
Working Paper No.
FR 15-25

July 17, 2015

Product Market Peers and Relative Performance Evaluation

Sudarshan Jayaraman
University of Rochester - Simon Business School

Todd T. Milbourn
Washington University in Saint Louis - Olin Business School

Hojun Seo
Washington University in Saint Louis - John M. Olin Business School

Product Market Peers and Relative Performance Evaluation

Sudarshan Jayaraman
Simon Business School
University of Rochester
jayaraman@simon.rochester.edu

Todd Milbourn
Olin School of Business
Washington University in St. Louis
milbourn@wustl.edu

Hojun Seo
Olin School of Business
Washington University in St. Louis
hojun.seo@wustl.edu

July 17, 2015

Abstract

Relative Performance Evaluation (RPE) theory predicts that firms filter out common shocks (i.e., those affecting the firm and its peers) while evaluating CEO performance, and that the extent of filtering increases with the number of firms in the peer group. Despite the intuitive appeal of the theory, previous tests of RPE find weak and inconsistent evidence. We hypothesize that one reason for the mixed evidence is the inaccurate classification of peers. Rather than using static, pre-defined Standard Industry Classifications (SIC), we exploit recent advances in textual analysis and define peers based on firms' product descriptions in their 10-K filings (e.g., Hoberg and Phillips (2015)). This alternative classification not only captures common shocks to firms' product markets more effectively, but also tracks the evolving nature of these product markets, as 10-Ks are updated annually. Using product market peers, we find three pieces of evidence consistent with RPE – (i) firms on average filter out common shocks to performance measures, (ii) the extent of filtering increases with the number of peers, and (iii) firms completely filter out common shocks in the presence of a large number of peers. We are able to replicate the first finding but not the others using SIC codes. Overall, our results suggest that a key identification strategy to testing RPE theory lies in accurately defining the peer group.

JEL codes: M40; M41; G30; J33
Keywords: Product market peers, CEO compensation, Relative Performance Evaluation

We appreciate helpful comments from Dirk Black, Patricia Dechow, John Donovan, Richard Frankel, Radha Gopalan, Jared Jennings, Ranjani Krishinan, Xiumin Martin, Partha Mohanram, Ram Natarajan, and workshop participants at 2015 Conference on Convergence of Financial and Managerial Accounting Research and Wash U. We thank Gerard Hoberg and Gordon Phillips for providing product market peers data. All errors remain our responsibility.

1. Introduction

Relative Performance Evaluation (RPE) theory predicts that a principal evaluates an agent based on the agent's idiosyncratic performance after filtering out the effect of common external shocks on the agent's overall performance (Holmstrom 1982; Holmstrom and Milgrom 1987). In addition, the RPE theory predicts that the optimal contract filters out common external shocks completely if the number of peers in the same team is sufficiently large, which allows the principal to obtain a more precise estimate of common external shocks (Holmstrom 1982; Gibbons and Murphy 1990). Prior research attempting to find the evidence of RPE documents weak and inconsistent results (i.e., the RPE puzzle, see Prendergast (1999), Lambert (2001), and Frydman and Jenter (2013) for discussions). To explain inconsistent results, several studies attribute the lack of RPE to the management's rent-seeking behaviors (e.g., Bertrand and Mullainathan 2001; Garvey and Milbourn 2006) while other studies attempt to find cross-sectional variations in the use of RPE (e.g., Garvey and Milbourn 2003; Rajgopal, Shevlin, and Zamora 2006; Gopalan, Milbourn, and Song 2010). In our study, we attempt to document evidence consistent with the RPE theory by identifying RPE peers based on textual analysis of firms' products in 10-K filings (Hoberg and Phillips 2013). By doing so, we hope to shed additional light on the RPE puzzle by investigating a potential reason why prior research fails to find consistent evidence of RPE and to provide direct evidence consistent with optimal contracting theory.

The key idea of RPE in compensation contracts is that CEOs must be evaluated by the firm's idiosyncratic performance after removing the effect of common external shocks on the individual firm performance.[1] In this sense, identifying peer firms plays a central role in RPE

[1] Firm performance mainly consists of a firm-specific component (i.e., idiosyncratic performance) and non-firm-specific component. The firm-specific component of performance are likely to reflect firm-specific or CEO-specific efforts that makes firm performance different from their rivals while the non-firm-specific component of the firm's performance is influenced by common external shocks that affect all firms in the same industry or market in a

research because one can only estimate the effect of common external shocks by observing peer firms' performance (Holmstrom 1982; Gibbons and Murphy 1990; Carter, Ittner, and Zechman 2009). However, the theory gives little guidance as to the form of common external shocks in practice (Gibbons and Murphy 1990). Intuitively, common uncertainty parameter in the theoretical model is likely to represent common demand and supply shocks (i.e., seasonal and random variations in product demand or price changes in input factors) to which all firms operating in the same product markets are exposed. If so, firms with similar products are more likely to experience these external shocks in a similar fashion because they have similar factor input suppliers, meet in similar product markets, and share similar customers. Hence, by aggregating performance of peer firms operating in the same product markets, researchers can calculate an estimate of common external shocks if the aggregation eliminates the idiosyncratic component of peer firms' performance on average.

To identify RPE peers, most prior RPE studies use pre-defined industry classifications, mostly Standard Industry Classifications (SIC), under the implicit assumption that firms sharing the same industry code are facing common external shocks. However, another stream of research suggests that firms in the same industry code are significantly heterogeneous, suggesting that they misclassify firms (Clarke 1989; Kahle and Walkling 1996; Bhojraj et al. 2003; Dopuch et al. 2008; Brickley and Zimmerman 2010; Guenther and Rosman 1994; Hoberg and Phillips 2013).[2] In addition, pre-defined industry classifications rarely change over time and, for this reason, those industry classifications might not properly capture the evolving nature of the firm's product

similar fashion. According to the RPE theory, a CEO of the firm must be rewarded based only on the change in the firm-specific component of performance because the non-firm-specific component of performance is not under the CEO's control but increases risks in the performance measure, leading to inefficient compensation contracts.

[2] It is worth to note that pre-defined industry classification such as SIC or NAICS focus on whether firms' production processes are similar rather than whether firms produce similar products. For example, "NAICS will be erected on a production-oriented, or supply-based, conceptual framework. This means that producing units that use identical or similar production processes will be grouped together in NAICS." http://www.naics.com/info.htm

markets as the firm's product offerings change (Hoberg and Phillips 2013). Hence, if the pre-defined industry classifications fail to classify firms experiencing common demand and supply shocks, aggregating performance of all firms in the same industry classification is more likely to provide a noisy measure of common external shocks, leading to the failure of detecting consistent evidence of RPE.

To capture the aforementioned product market aspects in identifying RPE peers, we use Text-based Network Industry Classifications (TNIC) developed by Hoberg and Phillips (2013). TNIC are explicitly constructed based on the product similarities among firms, allowing us to investigate whether firms' product market peers constitute a relevant peer group in RPE tests.[3] Hoberg and Phillips (2013) specifically calculate product similarity scores of all possible pairs of firms in each year by parsing firms' product descriptions in annual 10-K filings. If the similarity score between a firm and its potential peer firm is above the pre-determined similarity threshold, the peer firm is identified as the firm's product market peer. Therefore, each firm has its own distinct set of product market peers in this classification scheme. In addition, the composition of TNIC varies over time because TNIC are based on firms' product descriptions in 10-K filings, which are updated annually. This reflects the changing nature of the firm's product markets as the firm's product strategies change (Hoberg and Phillips 2013).[4] Hence, we argue that this

[3] There are potentially several ways to define a firm's RPE peers instead of relying on pre-defined industry classifications. For instance, one can define RPE peers based on peer firms chosen by analysts (De Franco et al. 2012), peer firms co-searched by investors (Lee, Wang, and Ma 2014), or peer firms self-selected and disclosed by the firm itself (Gong et al. 2011; Lewellen 2013). These classification schemes might also capture the product market peers. However, unlike TNIC, the underlying criteria of selecting peer firms in those classifications is *ex ante* unobservable.

[4] Hoberg and Phillip (2013) show that this text-based classifications better explain differences in industry characteristics such as profitability, sales growth, and market risk across industry. They show that positive (negative) industry demand shocks lead to more (less) firms entering into those industries. They also show that TNIC better reflect competitors identified by managers. Using TNIC, several studies find that this classification scheme provides a new insight regarding a firm's product market peers. For example, Hoberg and Phillip (2010) show that M&A transactions are more likely between firms having similar product descriptions and long-term outcome such as profitability is better when the target and the acquirer have similar product descriptions ex ante,

classification scheme allows us to estimate an empirical counterpart of common external shocks reflecting demand and supply shocks in product markets. To obtain a measure of common external shocks, we form quartile portfolios based on size and book-to-market ratio within each focal firm's TNIC industry and define RPE peers as firms in the same quartile portfolio as the focal firm. Equal-weighted average stock returns of those RPE peers give us a measure of common external shock (Peer Return, hereafter).

Using 19,105 firm-year observations spanning from 1996 to 2011, we first investigate whether the effect of common external shocks on the own firm performance is removed in revising total CEO compensation. Consistent with the RPE theory, we find that changes in total CEO compensation is positively associated with the firm's own stock returns (i.e., positive pay-for-performance sensitivity) while it is negatively associated with peer returns on average. In this case, the absolute magnitude of coefficient on peer return is significantly less than that of the own firm stock return, suggesting that firms on average partially filter out common external shocks in rewarding CEOs (i.e., the weak-form evidence of RPE, e.g., Gibbons and Murphy 1990).

We next examine whether the extent of filtering out common external shocks increases with the number of product market peers, and with a sufficiently large number of peers, the effect of common external shocks is completely filtered out as predicted by Holmstrom (1982). To this end, we divide our sample into three subsamples based on the number of product market peers (Hoberg and Phillips 2013) and perform the RPE test in each subsample. Consistent with expectations, we find that filtering out common external shocks in revising total CEO compensation increases significantly and monotonically as we move from the subsample with the lowest number of product market peers (Low group) to the subsample with the greatest number of

possibly due to product market synergies. Another example is Foucault and Fresard (2014) showing that a firm's investment is sensitive to the stock returns of product market peers.

product market peers (High group). Furthermore, we find that the effect of common external shocks on the firm's performance is completely filtered out in revising total CEO compensation in the High group (i.e., the strong-form evidence of RPE).[5] To our best of knowledge, this prediction has not been explicitly tested in prior research. We also find that the pay-for-performance sensitivities do not vary statistically regardless of the number of product market peers. Taken together, the results support our argument that product market peers constitute a relevant peer groups in empirical RPE tests.

We perform several additional tests to further support our argument. First, we examine the RPE theory using pre-defined industry classifications (SIC and GICS). Consistent with prior RPE studies, we find the weak-form evidence of RPE but we do not find the strong-form evidence of RPE (e.g., Antle and Smith 1986; Gibbons and Murphy 1990; Aggarwal and Samwick 1999a 1999b; Garvey and Milbourn 2006; Rajgopal et al. 2006). In addition, we do not find evidence on RPE increasing with the number of peers in pre-defined industries.

Second, we exploit the time-varying nature of the product market peer groups reflected in TNIC to shed additional lights on the importance of considering product market peers in RPE tests. We find that current stock returns of firms that exited in the focal firm's product markets in past periods (i.e., past peers) or firms that will enter the focal firm's product markets in future periods (i.e., future peers) in general do not provide information concerning common external shocks while current stock returns of current peers do provide relevant information.[6] These results suggest TNIC

[5] Holmstrom (1982) specifically states that "we would expect that with many agents we would be able to achieve approximately the same solution as if there were no common uncertainty at all."

[6] Intuitively, for example, if a firm's product market peer in period *t-1* exits and enters a new product market in period *t*, then the firm's product market peer in period *t-1* (i.e., past peers) might not be a relevant RPE peer in period *t* for the firm because the exited peer is not operating in the same product market in period *t* (i.e., they are not exposed to the same demand and supply shocks in period *t*). Foucalut and Fresard (2014) use this method and show that the average stock returns of past peer are unrelated to the focal firm's investment while the average stock returns of present (future) peers are strongly (weakly) related the focal firm's investment. We will explain this method in detail in section 4.

incorporates a firm's changing nature of product strategies in its construction and classify firms operating in the same product markets in a timely manner.

Third, we perform cross-sectional tests to ensure that we can find consistent results documented in prior research based on other optimal contracting theory. Gopalan, Milbourn, and Song (2010) propose a model showing that compensation committees should not remove the effect of common external shocks when a firm's exposure to common external shocks is the CEO's choice (i.e., strategic flexibility). We find consistent evidence that in situations where firms are more likely to incentivize CEOs to have strategic flexibility, RPE is less likely used in revising total CEO compensation.

Our study contributes to the existing compensation literature in two key ways. First, we find strong evidence of RPE in compensation contracts and highlight the fact that identifying peer firms that operate in the same product markets is an important aspect in empirical RPE research. In this sense, our study builds on the recent literature arguing that prior RPE research has failed to find the empirical evidence of RPE due to the incorrectly identified RPE peers (e.g., Albuquerque 2009; Gong et al. 2011; Lewellen 2013).

Second, most importantly, we document evidence supporting Holmstrom's prediction that RPE increases with the number of peers in the same team and with a sufficiently large number of peers, the solution is approximately equal to the case in which there are no common external shocks. Prior research attempts to find evidence supporting this prediction by examining whether product market competition is positively associated with RPE and document mixed evidence (Aggarwal and Samwick 1999; DeFond and Park 1997; Ali, Klasa and Yeung 2009).[7] We provide

[7] Aggarwal and Samwick (1999) find a negative association between product market competition and RPE in compensation contracts. DeFond and Park (1997) find a positive association between product market competition and RPE in CEO turnover decisions, but Ali, Klasa, and Yeung (2009) fail to replicate DeFond and Park (1997). Ali et al. (2009) point out that the product market competition measure used in DeFond and Park (i.e., Sales-based HHI)

direct evidence consistent with RPE increasing with the number of peers operating in the same product markets and a firm is optimally using RPE if the firm has a sufficiently large number of peers.

This paper proceeds as follows. In the next section, we discuss relevant literature and develop hypotheses. Next, we discuss empirical specifications to test RPE theory in CEO compensation contracts, and then we present the estimation results including robustness checks. Lastly, we conclude and summarize.

2. Relevant Literature and Hypothesis development

2.1. Literature Review

Holmstrom (1979) predicts that when the agent's efforts are unobservable and non-contractible, the second-best contracting mechanism is to provide an incentive contract that the payout is contingent upon performance of the agent in order to align the agent's interest with the principal's interest. Consistent with this prediction, prior research shows that revisions in CEO compensation are positively associated with the own firm stock returns (i.e., positive pay-for-performance sensitivity; Jensen and Murphy 1999; Aggarwal and Samwick 1999). This incentive contract, however, imposes additional risks on the agent due to external shocks that are not under control of the agent but affect the agent's individual outcomes. These uncontrollable shocks potentially decrease the utility of the agent because the agent is assumed to be risk averse, reducing contracting efficiency. Holmstrom (1982) predicts that if the principal can identify and filter out external shocks in evaluating performance of the agent, then contracting efficiency can be

is only based on sales of publicly-traded firms within the industry, resulting in the biased measure of product market competition. Hence, it is a still open question whether or not product market competition is positively associated with RPE in both CEO compensation contracts and CEO turnover decisions. Bushman and Smith (2001) call for research to resolve conflicting results in Aggarwal and Samwick (1999b) and DeFond and Park (1997).

improved.[8] That is, the agent is not rewarded simply for its own performance rather for the agent's idiosyncratic performance measured relative to its peers' performance.

Prior studies have attempted to find empirical evidence of this RPE theory in CEO compensation contracts (e.g., Antle and Smith 1986; Gibbons and Murphy 1990; Jensen and Murphy 1990; Janakiraman, Lambert, and Larcker 1992; Aggarwal and Samwick 1999a among others). However, the evidence in prior studies is at best mixed (Prendergast 1999; Lambert 2001; Frydman and Jenter 2013).

This lack of empirical evidence on RPE have motivated research investigating why there is a lack of consistent evidence of RPE. For example, Bertrand and Mullainathan (2001) argue the lack of RPE is attributed to the rent-seeking behavior of managers. They argue that firms with weak corporate governance are less likely to use RPE because CEOs in these firms can affect their pay-setting process and are paid for positive external shocks but not paid for negative external shocks (i.e., pay-for-luck literature). Other stream of research attempts to find cross-sectional variation of RPE by identifying factors that alter the costs and benefits of using RPE. For instance, Gopalan et al. (2010) show that if the exposure to the common external shocks depends on the CEO's choice to maximize firm value (i.e., strategic flexibility), then RPE is less likely to be used in compensation contracts to incentivize managers to take actions increasing firm value.

Another stream of research has attempted to explain the lack of empirical evidence of RPE by refining the empirical methods or concepts of peer firms in RPE tests. For instance, Albuquerque (2009) argues that using all firms in the same industry (SIC code) as RPE peers can

[8] Holmstrom (1982) specifically assumes *homogeneous* agents in the same team, and each agent's performance (x_i) is determined by common uncertainty parameter (η), which affects all agents in the same team, and idiosyncratic error term, which is determined by the agent-specific efforts (i.e., $x_i = \eta + e_i$). By aggregating performance of all agents in the same team, the idiosyncratic error terms are averaged out in the aggregate performance index, and thus the measure of common uncertainty parameter can be estimated.

be problematic because all firms in the same industry may not face common external shocks and firms' abilities to respond to common external shocks varies substantially even in the same industry. Albuquerque defines RPE peers as firms in the same size quartile portfolio in the same two-digit SIC industry and finds stronger evidence of RPE. Dikolli et al. (2011) analytically show that aggregating heterogeneous firm performance in the same industry adds significant summarization bias in the measure of common external shocks, leading to the failure in detecting RPE use.

All aforementioned studies assume that firms in the same industry are RPE peer firms. Several studies, however, point out that firms can self-select their peer firms that might be different from firms that researchers assume as RPE peers. Therefore, those studies rely on surveyed or disclosed peer firms to test the RPE theory (Murphy 1999; Bannister and Newman 2003; Carter et al. 2009; Gong et al. 2011; Lewellen 2013).[9,10] This line of research also argues that defining all firms in the same industry might lead to noisy measure of common external shocks, resulting in the failure of detecting empirical evidence of RPE in prior studies.

2.2 Product market peers

Another potential reason why prior studies relying on the pre-defined industry classification have failed to find the evidence of RPE might be that firms in the pre-defined

[9] For example, Gong et al. (2011) use compensation disclosures mandated by SEC after 2006 and examine the RPE theory. Interestingly, Gong et al. (2011) do not find evidence of RPE following the method used in Albuquerque (2009), but find evidence of weak-form evidence of RPE using self-selected RPE peers by the firm. Similar to Gong et al. (2011), Lewellen (2013) collects a firm's significant competitors disclosed in the firm's 10-K filings, and find that stronger evidence of RPE.

[10] The key difference of our study is that we directly test Holmstrom's prediction that RPE is increasing with the number of peers while those studies do not. Research using explicitly disclosed peer firms generally focuses on whether firms actually use RPE and the form of RPE in compensation contracts (e.g., tournament style of RPE). In addition, research using explicitly disclosed peer firms does not speak to whether or not firms use RPE through corporate boards' subjective discretion without disclosures (Ferri 2009) while findings in our study suggest that corporate boards are doing so.

industry are heterogeneous, resulting in the misclassification of firms (Clarke 1989; Kahle and Walkling 1996; Bhojraj et al. 2003; Dopuch et al. 2008; Hoberg and Phillips 2013). A potential reason for this misclassification could be that SICs and NAICS group firms based on their production functions rather than similar products or outputs (Bhojraj et al. 2003).[11]Guenther and Rosman (1994) also show that different databases (e.g., Compustat and CRSP) often assign different SIC codes for the same firm, which further suggests that firms in the same industry code might be heterogeneous. In addition, pre-defined industry classifications rarely change over time and thus, those classifications may not properly capture the evolving nature of the firm's product markets as the firm's product offerings change (Hoberg and Phillips 2013).

The RPE theory assumes homogenous agents in the same team that share the same common uncertainty parameter (Holmstrom 1982). Arguably, the empirical counterpart of common uncertainty parameter in practice likely represent common demand and supply shocks that affect all firms producing similar products, rather than, for example, having similar production functions. The distinction is important because having similar production functions does not necessarily imply those firms producing similar products (e.g., Bernard and Skinner 1996; Brickley and Zimmerman 2010). In addition, a firm enters and exit product markets if the firm starts or stops producing their products according to their product strategies change. In this case, the firm's product market peers also change accordingly. Traditional industry classifications, however, are not changed rapidly over time, which might lead to the failure in capturing this dynamic nature of evolving product markets. Therefore, if pre-determined industry classifications are not properly capturing firms' product market aspects, then RPE tests relying on pre-defined industry

[11] http://www.naics.com/info.htm. The Census Department states "NAICS was developed to classify units according to their production function. NAICS results in industries that group units undertaking similar activities using similar resources but does not necessarily group all similar products or outputs."

classifications might fail to detect consistent evidence. Regarding this point, Gibbons and Murphy (1990) state that "… [omitted by authors], although our inability to detect an industry effect after controlling for market movements may reflect the inappropriateness of industry definitions based on SIC codes for purposes of relative performance evaluation (p.49)."

Hence, we argue that the empirical analysis should identify peer firms producing similar products who face similar demand and supply shocks as RPE peers. To this end, we use Text-based Network Industry Classifictions (TNIC) developed by Hoberg and Phillips (2013) to identify RPE peers. Hoberg and Phillips (2013) identify peer firms based on the pairwise product similarity scores among firms by parsing firms' product descriptions in annual 10-K filings (Item 1 or 1A). They argue that firms producing similar products are more likely to be peer firms competing in the same product markets.

Hoberg and Phillips (2013) specifically convert each firm's product description in 10-K filings into a word vector and calculate product cosine similarity scores for every pair of firms (i.e., the distance between two word vectors for every pair of firms). For example, a firm *i*'s product similarity score with a firm *j* is calculated as the dot product of the word vector of the firm *i*, which consists of vocabularies describing the firm *i*'s products, and that of the firm *j*. This cosine product similarity score between firm *i* and firm *j* is bounded in [0,1] and increases with the number of same words that both firm *i* and firm *j* use, implying that firm pairs with high cosine similarity scores are likely to operate in the similar product markets. Firm *j* is classified as firm *i*'s product market peer if product similarity score between firm *i* and firm *j* is above a pre-specified minimum similarity threshold.[12] This classification yields a group of product market peers for every firm,

[12] Hoberg and Phillips (2013) state that "Although one can use any minimum similarity threshold to construct a classification, we focus on thresholds generating industries with the same fraction of membership pairs as SIC-3 industries, allowing us to compare our industries to SIC-3 in an unbiased fashion."

which allows peer group composition to vary year-to-year and firm-by-firm. Hoberg and Phillips (2013) argue that this procedure can capture the notion that the most appropriate peer firms are firms producing similar products. In addition, Hoberg and Phillips (2013) also argue that TNIC captures the changing nature of product markets over time because all firms' update their product descriptions annually and the updates are required to be correct and timely by SEC. Hence, we test the RPE theory using TNIC to examine whether or not firms' product market peers constitute relevant RPE peer groups.

2.3. The implication of the number of product market peers in RPE research

Holmstrom (1982) predicts that if the number of agents is sufficiently large enough to infer the precise value of common uncertainty parameter, the principal can completely filter out common uncertainty in evaluating an individual agent's performance. If the number of agents in a team is small, then idiosyncratic performance of agents is not sufficiently eliminated, resulting in the principal partially filtering out common uncertainty in evaluating the agent's performance (Gibbons and Murphy, 1990). In this sense, we anticipate that RPE in compensation contracts increases with the number of product market peers. In addition, we expect that common external shocks are completely filtered out if the firm has a sufficiently large number of product market peers. Hence, we state our main hypotheses below.

H1: Relative Performance Evaluation increases with the number of product market peers

H2: Firms completely filter out the effect of common external shocks in revising total compensation when firms have a large number of product market peers.

3. Research Design

3.1. Empirical specification

We use the following empirical specification, which is originally proposed by Holmstrom and Milgrom (1987) and has been widely used in prior RPE studies (e.g., Gibbons and Murphy, 1990).

$$\begin{aligned}\Delta \ln(\textit{Total Comp}) = {} & \alpha + \beta_1 \textit{Firm Return} + \beta_2 \textit{Peer Return} + \beta_3 \Delta \textit{Size} + \beta_4 \Delta \textit{Book-to-Market} \\ & + \beta_5 \textit{Idiosyncratic Volatility} + \beta_6 \ln(\textit{CEO Tenure}) + \beta_7 \textit{Ownership} \\ & + \Sigma \textit{Industry-Year Dummies} + \varepsilon\end{aligned}$$

The Δln(*Total compensation*) variable is the change in total CEO compensation, which is measured as the sum of salary, bonus, grant-date fair value of stock and option grants, long term incentive payouts, other annual compensation, and all other annual compensation (TDC1 in ExecuComp). The *Firm Return* variable captures firm *i*'s own performance and measured by firm *i*'s annual buy-and-hold stock returns including dividends. The *Peer Return* variable captures the amount of common external shocks and measured by the equal-weighted annual stock returns of product market peers excluding firm *i*. To define product market peers, we construct quartile portfolios within each firm's TNIC group based on the market value of equity (i.e., Size) and book-to-market ratio at the beginning of the period in each year (e.g., Daniel and Titman 1997). Firms excluding firm *i* in the same quartile portfolio as firm *i* are defined as firm *i*'s product market peers in period *t*. The equation is intended to examine whether the coefficient on the *Firm Return* variable is statistically positive, which captures CEOs rewarded for taking actions increasing shareholder wealth (i.e., pay-for-performance sensitivity), while the *Peer Return* variable is statistically negative, which captures common external shocks filtered out in evaluating CEOs' efforts. In addition, optimal contracting theory predicts that the sum of the coefficients on the *Firm*

Return and the *Peer Return* variables is statistically zero if common external shocks are completely removed from the own firm performance, indicating that CEOs are evaluated solely based on the idiosyncratic performance of the firm (Holmstrom and Milgrom 1987).

Following prior studies, we also include control variables in the equation. We include changes in total assets and changes in book-to-market ratio to control size and growth option of the firm (Smith and Watts 1992). We include idiosyncratic volatility (Aggarwal and Samwick 1999), the natural logarithm of CEO tenure, and CEO ownership to control the effect of firm and CEO characteristics on firms' compensation policies. We also include industry-year fixed effects to control unobservable time-varying industry specific factors (Gormley and Matsa 2014).[13]

4. Data and Descriptive Statistics

We retrieve market values of equity and stock returns data from CRSP, financial statement data from COMPUSTAT, and CEO compensation data from ExecuComp. Following Garvey and Milbourn (2006), we use a sample of ExecuComp firms with at least two consecutive years of data for each CEO during our sample period. We delete observations with missing financial and compensation data. The above data requirements yield a sample of 19,105 firm-CEO-year observations. The sample period ranges from 1996 to 2011 because TNIC are available during this sample period (Hoberg and Phillips 2013).

Table 1 presents descriptive statistics. The mean (median) of total compensation is $5.2 million ($2.83 million), which shows significant right skewness as in prior compensation studies (e.g, Albuquerque, 2009). Hence, we take the natural logarithm in total compensation to reduce skewness. Table 2 presents correlations among main variables. The mean (median) value of

[13] For industry fixed effects, we use Fixed Industry Classifications that is also constructed based on the product similarity scores (Hoberg and Phillips, 2013) to be consistent with our use of TNIC in our main analyses.

correlations between the *Firm Return* variable and the *Peer Return* variable based on TNIC, SIC, and GICS are 0.56, 0.93, and 0.57, respectively. These correlations suggest that own firm performance is significantly correlated with its peer firms' performance. We also note that changes in total compensation are positively correlated with most performance measures. Changes in total compensation are also positively correlated with changes in firm size (0.15) and negatively correlated with book-to-market ratio (-0.08), suggesting that large and growth firms incur greater compensation costs to hire high-ability managers (Smith and Watts 1992).

5. Empirical results

Table 3 presents the result of estimating equation (1). In column (1) of Table 3, we demonstrate the estimation result using the full sample. Consistent with prior RPE literature, we find significantly positive coefficient on the *Firm Return* variable at 1% level and significantly negative coefficient on the *Peer Return* variable at 1% level, but the absolute value of the coefficient on the *Peer Return* variable is significantly less than the absolute value of the coefficient on the *Firm Return* variable (F-statistics of 41.57). This result suggests that CEOs are rewarded by positive stock returns and common external shocks are on average partially filtered out in evaluating CEOs' efforts (i.e., the weak-form evidence of RPE). Said differently, optimal contracts on average partially incorporate the estimate of common external shocks because the *Peer Return* variable measures the common external shocks with error (Gibbons and Murphy 1990).

Next, we examine whether or not the extent of RPE usage increases as the number of product market peers increases (Holmstrom 1982). To this end, we divide the full sample into three subsamples based on the number of product market peers (Hoberg and Phillips 2013) and estimate

the equation (1) for each subsample. Specifically, firm-year observations that belong to the first, second, and third tercile in terms of the number of product market peers are classified as the Low, Medium, and High group, respectively. We expect to observe negative coefficients on the *Peer Return* variable decreasing significantly as we move from the Low group to the High group. Consistent with our expectations, we find negative but insignificant coefficient on the *Peer Return* variable in the Low group in column (2) but the coefficient decrease monotonically and significantly as we move from the Low group to the Medium group in column (3) or to the High group in column (4). We test the coefficients' statistical difference and find that the coefficient on the *Peer Return* variable in the High group is significantly different from the coefficient on the *Peer Return* variable in the Low group at 1% level (p-Value 0.007). We note that the coefficient on the *Firm Return* variable does not vary significantly across all subsamples. The difference of coefficients on the *Firm Return* variable between the Low group and the High group is not statistically significant (p-value 0.928). These results suggest that the number of product market peers is associated with RPE but not associated with the incentive contracts itself.

Lastly, we examine whether firms with a sufficiently large number of product market peers completely filter out common external shocks in revising total CEO compensation. In column (4) of Table 3, we find that the sum of the coefficient on the *Firm Return*, variable and the coefficient on the *Peer Return* variable is statistically zero (F-Statistics of 0.05), suggesting that common external shocks are completely filtered out in revising total CEO compensation. This evidence is consistent with Holmstrom (1982)'s prediction, which is also known as the strong-form evidence of RPE. Overall, results documented in Table 3 suggest that firms use RPE in rewarding their CEOs, and the extent of RPE usages hinges on the presence of enough number of product market peers.

6. Additional tests

6.1. Pre-defined industry classifications to identify RPE peers

In our empirical analyses, we use TNIC to identify RPE peers. In this section, we replicate our results using pre-defined industry classifications. In Panel A of Table 4, we use three-digit SIC codes and calculate the *Peer Return* (*SIC*) variable based on the same method described above.[14] In column (1) of Table 4, we find the same result documented in column (1) of Table 3 by using SIC industries to identify RPE peers (i.e., the weak-form evidence of RPE). We note that this result is consistent with prior RPE research (Albuquerque 2009). Next, we partition the full sample into three subsamples based on the number of firms in the same SIC industry and estimate equation (1) in each subsample. Contrary to the results in Table 3, we do not find evidence consistent with our hypothesis that RPE increases with the number of firms in the same SIC industry. We also do not find evidence that firms with a number of SIC industry peers are completely filtering out common external shocks. In Panel B of Table 4, we use Global Industry Classification Standard (GICS) codes to define the firm's RPE peers. GICS codes are the most recent and improved classification method developed by MSCI Inc. and Standar&Poor's. Bhojraj et al. (2003) document that firms in the same GICS classifications have higher profitability and growth correlations than firms that share the SIC codes, NAICS codes, and Fama-French classification codes. They conclude that GICS is a better industry classification to identify industry peers that compete in the same product markets. Using GICS to define RPE peers, we perform same analyses and find that firms on average partially filter out common external shocks in the full sample. However, once again, we

[14] TNIC is comparable with three-digit SIC because the pre-specified minimum product similarity threshold use in constructing TNIC is set to generate industries with the same fraction of industry pairs as three-digit SIC industries (Hoberg and Phillips, 2013). Results using NAICS are similar and are not tabulated.

do not find evidence consistent with the theory that the number of firms in the same GICS industry is associated with the extent to which firms use RPE in compensation contracts.

To investigate the potential reasons why we do not find consistent evidence using pre-defined industry classifications, we include both the *Peer Return* variable based on TNIC and the *Peer Return* variable based on the pre-defined industry classification simultaneously in the same regression analysis. Table 5 demonstrates the results. We find that only the coefficient on the *Peer Return* variable based on TNIC is significantly negative at 5% level while the coefficient on the *Peer Return* (*SIC*) in column (1) and the coefficient on the *Peer Return* (*GICS*) variable in column (2) are not statistically significant. These results suggest that proxies of common external shocks based on pre-defined industry classifications capture some portion of demand and supply shocks in product markets but there are still remaining portion of common external shocks that is not captured by these proxies but is filtered out in compensation contracts.

6.2. Dynamic peer groupings

As noted earlier, one of the key advantage of using TNIC to identify RPE peers is that TNIC captures the changing and evolving nature of the firm's product markets. Therefore, we can examine whether or not current stock returns of past, current, and future product market peers contain information regarding common external shocks. For example, consider *past* peer firm *j* that was the product market peer of firm *i* in period *t-1* but not in period *t* (i.e., firm *j* exited firm *i*'s product space in period *t-1*). In this case, firm *j*'s current stock return in period *t* is less likely to contain information regarding common demand and supply shocks that firm *i* faces in period *t* because firm *i* and firm *j* are not operating in the same product market in period *t*. Similarly, if firm *j* is not a product market peer of firm *i* but is expected to be a peer in period *t+1* (i.e., future

peer), then the current stock returns of firm *j* in period *t* is also less likely to contain relevant information regarding common external shocks that firm *i* is experiencing in period *t*. However, if entering new product markets takes time, then firm *j* is likely taking some activities to enter the new product market in period *t* (e.g., investments), resulting in firm *j*'s stock returns in period *t* might contain information regarding common external shocks (Foucault and Fresard, 2014). Foucault and Fresard (2014) first adopt this approach and show that past (future) peers' stock price is not (weakly) associated with the focal firm's investment while present peers' stock price is informative to the focal firm's investment.

Similar to Foucault and Fresard (2014), we classify firm-year observations into five sets of peer firms: (1) past peers, (2) expected to exit peers, (3) current peers, (4) new peers, and (5) future peers. We define *past peers* as firms that were in the same TNIC group as the focal firm in period *t-1* but not in period *t* any more. We define *expected to exit peers* as firms that are in the same TNIC group as the focal firm in period *t* but will not be in period *t+1*. *Current peers* are defined as firms that are in the same TNIC group as the focal firm in period *t-1* as well as in period *t*. *New peers* are firms that were not in the same TNIC group in period *t-1* but are in the same TNIC group in period *t*. Lastly, we define *future peers* as firms that are not in the same TNIC group as the focal firm in period *t* but will be in the same TNIC group in period *t+1* (or *t+2*). We then calculate equal-weighted stock returns of each set of peers using stock returns in period *t* and replace these the *Peer Return* variable in equation (1) with each of these calculated stock returns.

Table 6 reports the estimation results. Consistent with our expectations, in column (1) of Panel A, the *Peer Return* variable based on past peers is not associated with changes in total CEO compensation, suggesting that current stock returns of past peers (i.e., firms exited the product market of the focal firm in period *t-1*) do not contain information regarding common external

shocks for the focal firm in period *t* possibly because past peers are not operating in the same product market as the focal firm in the current period.

In column (2), (3), and (4) of Panel A, we find that the coefficients on the *Peer Return* variable based on the expected to exit peers, current peers, and new peers are significantly negative at 1% level (i.e., all of them are peers in period *t*) but the magnitude of the coefficient on the *Peer Return* variable based on current peers is greater than those of expected to exit peers and new peers. In Panel B, we compare the statistical differences between those three variables by including two of them simultaneously in the same regression. In column (1) of Panel B, we include the *Peer Return* variables based on current peers and new peers simultaneously, and find that both coefficients are statistically negative at 10% level. This result suggests that both variables contain relevant and distinct information regarding common external shocks. In column (2) and (3) of Panel B, we include the *Peer Return* variables based on expected to exit peers and current peers in column (2) or new peers in column (3), respectively, and find that the coefficient on the *Peer Return* variable based on expected exit peer is not statistically significant while the coefficient on the *Peer Return* variables based on either current peers or new peers are statistically negative. These results suggest that even though stock returns of expected to exit peers contain relevant information concerning common external shocks, this is subsumed by other variables.

Back to column (5) of Panel A, we find that the coefficients on the *Peer Return* variable based on future peers at period *t+1* are significantly negative at 1% level. As noted earlier, this relation is expected if entering new product markets takes time, resulting in a lag between period t and period *t+1* (Foucalut and Fresard, 2014). Therefore, we further identify future peers at period *t+2*, which are defined as firms that are not in the same TNIC group as the focal firm in period *t* but will be in the same TNIC group in period *t+2*, and calculate equal-weighted average stock

returns for those firms. In column (6) of Panel B, we use the calculated average stock returns using future peers at period *t+2* in the regression and find that the coefficient of the *Peer Return* variable based on future peers at period *t+2* is not associated with changes in CEO total compensation. Similarly, we compare the statistical differences between the *Peer Return* variable based on new peers and that of future peers at period *t+1* or period *t+2* in column (4) and (5) of Panel B, respectively. We find that the coefficients on the *Peer Return* variable based on future peers at period *t*+1 or period *t+2* in column (4) and column (5), respectively, are not significant while the coefficients on the *Peer Return* variables based on new peers are statistically negative. Overall, the results in Table 5 further corroborate our argument that a firm's product market peers have an important implications in RPE tests.

6.3. Strategic flexibility and RPE

Next, we investigate cross-sectional variations in RPE based on the other optimal incentive contracting theory. Gopalan, Milbourn, and Song (2010) propose a model showing that the use of RPE decreases if firms want to provide strategic flexibility to their CEOs. Gopalan et al. argue that "the board of directors is not primarily concerned with how hard the CEO is actually working, but whether she has the vision to choose the right strategy for deploying the firm's assets. In doing so, the CEO's concern is with the firm's strategic direction in lieu of its surrounding market environment." Put differently, if the effect of common external shocks on firm performance is not random but under the CEO's control, then the effect of common external shocks should not be excluded in evaluating the CEO's efforts.

Following Gopalan et al., we use two alternative proxies to identify firms that offer greater strategic flexibility to the CEO. First, we use the firm's market-to-book ratio. Firms with high

market-to-book ratios are more likely to have greater growth options and thus are more likely to provide their CEOs with greater strategic flexibility to allow more discretions in exercising those options. We classify firm-years with market-to-book ratios above the median as offering greater strategic flexibility to the CEO. Second, we use the peer-adjusted stock returns during the previous year as a measure of CEO talent. Gopalan et al. predicts that RPE is reduced for more talented CEOs due to the decreasing disutility of effort for more talented CEOs. We classify firm-years with the peer-adjusted stock returns during period *t-1* above the median as having more talented CEOs because firms managed by more talented CEOs are more likely to exhibit better peer-adjusted stock performance.[15]

Gopalan et al. further argue that if less RPE allows CEOs to have greater strategic flexibility, we expect to observe some evidence that CEOs with less RPE exploit the strategic flexibility to a greater extent at the firm level. Hence, following Gopalan et al., we identify firms with asset growth in subsequent period as exploiting their strategic flexibility to a greater extent and examine whether firms with asset growth (decline) in period *t+1* are less (more) likely to use RPE in period *t*.

Table 7 demonstrates the estimation results. In column (1) and (2), we divide the full sample into two subsamples based on the median of market-to-book ratios. We find that the coefficient on the *Peer Return* variable in column (2) (i.e., more growth options) is statistically greater at 5% level relative to the coefficient in column (1) (i.e., less growth options). In column (3) and (4), we use the peer-adjusted stock returns as a measure of the CEO talent, and find similar

[15] Gopalan et al. also examine whether multi-segment firms (based on the SIC industry) are less likely to use RPE. We do not examine this variable because segment information in Compustat is only based on SIC industries, which we do not rely on in our study. Gopalan et al. also use R&D expenditures to test the theory. In untabulated tests, we find that the extent of RPE in firms with high R&D expenditures is not significantly different from RPE in firms with low R&D expenditures. This result could be attributed to R&D expenditures being a noisier measure of the firm's growth options because a significant portion of firms in Compustat universe does not report R&D expenditures separately.

results. The coefficient on the *Peer Return* variable is statistically greater at 10% level in column (4) (i.e., greater CEO talent) relative to the coefficient in column (3) (i.e., lower CEO talent). In column (5) and (6), we use the asset growth rate in period *t+1* to investigate whether CEOs with less RPE exploit their strategic flexibility in subsequent period to a greater extent at the firm level. Consistent with our expectations, we find that firms with asset growth in period *t+1* filter out common external shock in a lesser extent in period *t* (i.e., less RPE) as evidenced by the significantly greater coefficient on the *Peer Return* variable in column (6) (i.e., asset growth in period *t+1*) relative that in column (7) (i.e., asset decline in period *t+1*). In sum, the results in Table 7 are in general consistent with Gopalan et al. and strengthen our argument that product market peers provide a valid measure for common external shocks.

7. Conclusion

This study attempts to find evidence supporting the theory of Relative Performance Evaluation (Holmstrom 1982) using product market peers identified by textual analysis of firms' product descriptions in 10-K filings (Hoberg and Phillip 2013). Holmstrom (1982) predicts that a principal filters out common external shocks to insulate the agent, resulting in improved contracting efficiency. In addition, Holmstrom (1982) predicts that the extent of filtering out increases with the number of peers in the same team. Prior RPE research, however, generally finds mixed and weak empirical evidence on this theory. We identify RPE peers based on firms operating in similar product markets (Hoberg and Phillips 2013), and find evidence consistent with the RPE theory. Furthermore, we find that the extent to which firms filter out common external shocks increases with the number of product market peers as predicted by the RPE theory. Based on these results, we argue that considering firms' product markets is an important factor in

empirical RPE research. This argument is further strengthened by additional robustness tests. Overall, our study highlight the importance of peer group composition in RPE tests that prior RPE studies have generally ignored and find strong evidence consistent with optimal contracting theory.

References

Aggarwal, R., Samwick, A., 1999a. The Other Side of the Trade-Off: The Impact of Risk on Executive Compensation. *Journal of Political Economy* 107, no. 1: 65–105.

Aggarwal, R., Samwick, A., 1999b. Executive Compensation, Strategic Competition, and Relative Performance Evaluation: Theory and Evidence. *The Journal of Finance* 54, no. 6 (1999).

Albuquerque, A., 2009. Peer Firms in Relative Performance Evaluation. *Journal of Accounting and Economics* 48, no. 1: 69–89.

Antle, R., Smith, A., 1986. An Empirical Investigation of the Relative Performance Evaluation of Corporate Executives. *Journal of Accounting Research* 24, no. 1: 1–39.

Bannister, J., Newman, H., 2003. Analysis of Corporate Disclosures on Relative Performance Evaluation. *Accounting Horizons* 17, no. 3: 235–246.

Bernard, V., Skinner, D., 1996. What Motivates Managers' Choice of Discretionary Accruals? *Journal of Accounting and Economics* 22, no. 1: 313–325.

Bertrand, M., Mullainathan, S., 2001. Are CEOs Rewarded for Luck? The Ones Without Principals Are. *The Quarterly Journal of Economics* 116, no. 3: 901–932.

Bhojraj, S., Lee, C., Oler, D., 2003. What's My Line? A Comparison of Industry Classification Schemes for Capital Market Research. *Journal of Accounting Research* 41, no. 5: 745–774.

Brickley, J., Zimmerman, J, 2010. Corporate Governance Myths: Comments on Armstrong, Guay, and Weber. *Journal of Accounting and Economics* 50, no. 2: 235–245.

Bushman, R., Smith, A., 2001. Financial Accounting Information and Corporate Governance. *Journal of Accounting and Economics* 32, no. 1: 237–333.

Carter, M., Ittner, D., Zechman, S., 2009. Explicit Relative Performance Evaluation in Performance-vested Equity Grants. *Review of Accounting Studies* 14, no. 2–3: 269–306.

Clarke, R., 1989. SICs as Delineators of Economic Markets. *Journal of Business*: 17–31.

Daniel, K., Grinblatt, M., Titman, S., and Wermers, R., 1997. "Measuring Mutual Fund Performance with Characteristic-based Benchmarks." *The Journal of Finance* 52, no. 3: 1035–1058.

DeFond, M., Park, C., 1999. The Effect of Competition on CEO Turnover. *Journal of Accounting and Economics* 27, no. 1: 35–56.

De Franco, G., Hope, O. K., and Larocque, S. 2012. Analysts' choice of peer companies. *Review of Accounting Studies*, 1-28.

Dikolli, S., Hofmann, C., Pfeiffer, T., 2011. Relative Performance Evaluation and Peer-performance Summarization Errors. *Review of Accounting Studies*: 1–32.

Dopuch, N., Mashruwala, R., Seethamraju, C., Zach. T, 2012. The Impact of a Heterogeneous Accrual-generating Process on Empirical Accrual Models. *Journal of Accounting, Auditing & Finance* 27, no. 3: 386–411.

Ferri, F., 2009. Discussion of 'Explicit Relative Performance Evaluation in Performance-vested Equity Grants'. *Review of Accounting Studies 14*, no. 2–3 (September 1): 307–313.

Foucault, T., and Fresard, L. 2014. Learning from peers' stock prices and corporate investment. *Journal of Financial Economics*, 111(3), 554-577.

Frydman, C., and Jenter, D. 2010. CEO Compensation. *Annual Review of Financial Economics*, *2*(1), 75-102.

Garvey, G., Milbourn, T., 2003. Incentive Compensation When Executives Can Hedge the Market: Evidence of Relative Performance Evaluation in the Cross Section. *The Journal of Finance* 58, no. 4: 1557–1582.

Garvey, G., Milbourn, T., 2006, Asymmetric Benchmarking in Compensation: Executives Are Rewarded for Good Luck but Not Penalized for Bad. *Journal of Financial Economics* 82, no. 1: 197–225.

Gibbons, R., Murphy, K., 1990. Relative Performance Evaluation for Chief Executive Officers. *Industrial and Labor Relations Review*: 30S–51S.

Gormley, T. A., and Matsa, D. A. 2014. Common errors: How to (and not to) control for unobserved heterogeneity. *Review of Financial Studies*, 27(2), 617-661.

Gong, G., Li, L. Y., Shin, J. Y., 2011. Relative Performance Evaluation and Related Peer Groups in Executive Compensation Contracts. *The Accounting Review* 86, no. 3 (2011): 1007–1043.

Gopalan, R., Milbourn, T., Song. F., 2010. Strategic Flexibility and the Optimality of Pay for Sector Performance. *Review of Financial Studies* 23, no. 5: 2060–2098.

Guenther, D., Rosman, A., 1994. Differences Between COMPUSTAT and CRSP SIC Codes and Related Effects on Research. *Journal of Accounting and Economics* 18, no. 1: 115–128.

Hoberg, G., Phillips, G., 2013. Text-Based Network Industries and Endogenous Product Differentiation. *Working Paper*.

Holmstrom, B., 1979. Moral Hazard and Observability. *The Bell Journal of Economics*: 74–91.

Holmstrom, B., 1982. Moral Hazard in Teams. *The Bell Journal of Economics*: 324–340.

Holmstrom, B., Milgrom, P., 1987. Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica: Journal of the Econometric Society*: 303–328.

Janakiraman, S., Lambert, R., Larcker, D., 1992. An Empirical Investigation of the Relative Performance Evaluation Hypothesis. *Journal of Accounting Research* 30, no. 1: 53–69.

Jensen, M., Murphy, K., 1990. Performance Pay and Top-Management Incentives. *The Journal of Political Economy* 98, no. 2: 225–264.

Kahle, K., Walkling, R., 1996. The Impact of Industry Classifications on Financial Research. *Journal of Financial and Quantitative Analysis* 31, no. 3.

Lambert, R., 2001. Contracting Theory and Accounting. *Journal of Accounting and Economics* 32, no. 1: 3–87.

Lee, C., Ma, P., Wang, C. 2014. "Search Based Peer Firms: Aggregating Investor Perceptions through Internet Co-searches." *Journal of Financial Economics*, forthcoming

Milbourn, T., 2003. CEO Reputation and Stock-based Compensation. *Journal of Financial Economics* 68, no. 2: 233–262.

Murphy, K., 1999. Executive Compensation. *Handbook of Labor Economics* 3: 2485–2563.

Prendergast, C., 1999. The Provision of Incentives in Firms. *Journal of Economic Literature* 37, no. 1: 7–63.

Rajgopal, S., Shevlin, T., Zamora, V., 2006. CEOs' Outside Employment Opportunities and the Lack of Relative Performance Evaluation in Compensation Contracts. *The Journal of Finance* 61, no. 4: 1813–1844.

Appendix. Variable Definitions

Variable	Definition
Total Compensation	*Total Compensation* is *TDC1* in ExecuComp, which is measured by the sum of salary, bonus, long-term incentive payouts, fair value of stock and option grants, and all other compensation for firm *i* in period *t.*
Δ*Total Compensation*	Δ*Total Compensation* is measured as changes in total compensation for firm *i* in period *t.*
ln(Δ*Total Compensation*)	*ln*(Δ*Total Compensation*) is measured as the natural logarithm of changes in total compensation for firm *i* in period *t.*
Firm Return	*Firm Return* is measured as the natural logarithm of one plus firm *i*'s annual buy-and-hold stock return in period *t*.
Peer Return	*Peer Return* is measured as the natural logarithm of one plus equal-weighted annual returns of firm *i*'s product market peers in period *t*. To define product market peers, we form quartile portfolios based on market value of equity and book-to-market ratio at the beginning of the fiscal year within the same TNIC group as firm *i* in period *t*. Product market peers are defined as firms that belong to the same quartile portfolio as firm *i* in period *t*. We require firm *i* to have minimum two peer firms in each period.
Idiosyncratic Volatility	*Idiosyncratic Volatility* is measured as the annualized standard deviations of residuals from the regression of firm *i*'s monthly returns on monthly equal-weighted peer returns using past 36 months (a minimum of 12 monthly observations is required).
# of Peers	*# of Peers* is measured as the number of product market peers for firm *i* in period *t*.
Δ*Size*	Δ*Size* is measured as changes to the natural logarithm of firm *i*'s total assets in period *t*.
Δ*Book-to-Market*	Δ*Book-to-Market* is measured as changes to firm *i*'s Book-to-Market ratios in period *t.* Book-to-Market is measured as book value of equity divided by market value of equity. Book value of equity is measured by shareholders' equity plus deferred tax and investment credit minus preferred stock. Market value of equity is obtained from CRSP and is calculated by the number of common shares outstanding multiplied by share price at the beginning of period *t*.

CEO Tenure	*CEO Tenure* is defined as the difference between the *BECAMECEO* variable in ExecuComp and the date of fiscal year-end for firm *i* at the beginning of period *t*.
ln(*CEO Tenure*)	*ln*(*CEO Tenure*) is measured as the natural logarithm of CEO tenure for firm *i* at the beginning of period *t.*
Ownership	*Ownership* is calculated as the number of shares owned by CEO excluding option divided by the number of shares outstanding for firm *i* at the beginning of period *t.*

Table 1 Descriptive Statistics

This table reports descriptive statistics for all sample firms with available information. The sample period ranges from 1996 to 2011. All variables are defined in Appendix A and all continuous variables are winsorized at the 1 and 99th percentiles.

	N	Mean	Std	Q1	Median	Q3
Total Compensation	19,105	5,201	10,886	1,347	2,832	5,904
Δ*Total Compensation*	19,105	208	11,945	-535	126	1,030
ln(Δ*Total Compensation*)	19,105	0.06	0.67	-0.22	0.07	0.36
Firm Return	19,105	0.13	0.49	-0.17	0.08	0.33
Peer Return	19,105	0.12	0.36	-0.10	0.10	0.29
Peer Return (*SIC*)	18,698	0.11	0.37	-0.11	0.09	0.29
Peer Return (*GICS*)	18,620	0.11	0.35	-0.10	0.10	0.28
# of Peers	19,105	20	28	4	10	23
# of Peers (*SIC*)	18,698	21	28	3	7	27
# of Peers (*GICS*)	18,620	17	23	5	11	19
Δ*Size*	19,105	0.08	0.21	-0.01	0.08	0.17
Δ*Book-to-Market*	19,105	0.03	0.33	-0.09	0.01	0.12
Idiosyncratic Volatility	19,105	0.10	0.05	0.06	0.09	0.12
CEO Tenure	19,105	8.77	7.49	3.59	6.42	11.33
ln(*CEO Tenure*)	19,105	1.86	0.78	1.28	1.86	2.43
Ownership	19,105	0.02	0.05	0.00	0.00	0.01

Table 2 Pearson / Spearman Correlation

This table presents Pearson (Above) / Spearman (Below) correlations. Correlations that are significant at 1% level are bolded. The sample period ranges from 1996 to 2011. All variables are defined in Appendix A and all continuous variables are winsorized at the 1 and 99th percentiles.

	ln(Δ*Total Comp.*)	*Firm Return*	*Peer Return*	*Peer Return* (*SIC*)	*Peer Return* (*GICS*)	Δ*Size*	Δ*Book-to-Market*	*Idio. Volatility*	ln(*CEO Tenure*)	*Ownership*
ln(Δ*Total Comp.*)	-	**0.18**	**0.07**	**0.07**	**0.08**	**0.15**	**-0.08**	**-0.04**	0.03	0.00
Firm Return	**0.21**	-	**0.56**	**0.53**	**0.57**	**0.16**	**-0.60**	**0.06**	0.01	0.01
Peer Return	**0.08**	**0.56**	-	**0.72**	**0.76**	**0.05**	**-0.40**	**0.04**	0.01	0.01
Peer Return (*SIC*)	**0.08**	**0.54**	**0.73**	-	**0.77**	**0.04**	**-0.39**	**0.02**	0.01	0.00
Peer Return (*GICS*)	**0.09**	**0.57**	**0.77**	**0.78**	-	**0.04**	**-0.41**	**0.04**	0.01	0.00
Δ*Size*	**0.17**	**0.18**	**0.06**	**0.06**	**0.05**	-	0.03	-0.03	**0.06**	**0.03**
Δ*Book-to-Market*	**-0.13**	**-0.77**	**-0.46**	**-0.45**	**-0.47**	0.00	-	**-0.03**	0.01	0.01
Idio. Volatility	**-0.03**	**-0.07**	**-0.04**	**-0.05**	**-0.04**	-0.01	**0.02**	-	-0.01	**0.09**
ln(*CEO Tenure*)	0.01	0.01	0.01	0.00	0.00	**0.08**	0.02	0.01	-	**0.36**
Ownership	0.00	-0.01	-0.01	-0.02	-0.01	**0.06**	0.03	**0.23**	**0.45**	-

Table3 Tests of Relative Performance Evaluation

This table presents the results obtained from the regression of the natural logarithm of the change in total CEO compensation on the own firm stock returns (*Firm Return*), peer returns based on TNIC (*Peer Return*), control variables, and industry-year fixed effects during the sample period 1996-2011. All variables are winsorized at 1% and 99% levels and are defined in the Appendix A. In column (2), (3), and (4) the full sample is divided into three subsamples based on the number of firms in the same TNIC. Standard errors are clustered by firm. ***, **, and * represent significance level at the 1%, 5%, and 10% level, respectively. Robust t-statistics are in parentheses.

		Dependent variable: ln(Δ*Total Compensation*)			
Independent Variables		(1)	(2)	(3)	(4)
			# of TNIC Peers		
		Full Sample	Low	Med	High
Firm Return		0.331***	0.337***	0.355***	0.333***
		(16.380)	(10.196)	(8.570)	(8.663)
Peer Return		**-0.120*** **	**-0.055**	**-0.112** **	**-0.313*** **
		(-4.117)	**(-1.607)**	**(-1.995)**	**(-3.619)**
Δ*Size*		0.266***	0.322***	0.238***	0.235***
		(8.084)	(5.111)	(3.905)	(3.752)
Δ*Book-to-Market*		0.076***	0.096**	0.110**	0.025
		(3.077)	(2.488)	(2.273)	(0.564)
Idiosyncratic Volatility		0.020	0.380**	-0.090	-0.089
		(0.202)	(2.069)	(-0.390)	(-0.445)
ln(*CEO Tenure*)		0.017***	0.006	0.026**	0.019*
		(3.166)	(0.518)	(2.169)	(1.912)
Ownership		-0.045	-0.066	-0.178	0.386*
		(-0.574)	(-0.456)	(-1.110)	(1.850)
Constant		-0.002	-0.021	-0.010	0.010
		(-0.164)	(-0.732)	(-0.319)	(0.375)
Test strong RPE: $\beta_1 + \beta_2 = 0$	**F-Stat**	**41.57**	**42.04**	**14.59**	**0.05**
	p-value	**(0.00)**	**(0.00)**	**(0.00)**	**(0.82)**
Industry-Year Fixed Effects		Yes	Yes	Yes	Yes
# of observations		19,105	6,875	6,050	6,180
Adjusted R-squared		0.083	0.118	0.103	0.094

Coefficient difference (High - Low)	Coeff. Diff.	*p*-values
Firm Return	-0.005	(0.928)
Peer Return	-0.258***	(0.007)

Table 4 RPE tests based on pre-defined industry classifications

This table presents the results obtained from the regression of the natural logarithm of the change in total CEO compensation on the own firm stock returns (*Firm Return*), either peer returns based on SIC industries (*Peer Return* (SIC)) in Panel A or peer returns based on GICS industries (*Peer Return* (GICS)) in Panel B, control variables, and industry-year fixed effects during the sample period 1996-2011. All variables are winsorized at 1% and 99% levels and are defined in the Appendix A. In column (2), (3), and (4) the full sample is divided into three subsamples based on the number of firms in each industry classification. Standard errors are clustered by firm. ***, **, and * represent significance level at the 1%, 5%, and 10% level, respectively. Robust t-statistics are in parentheses.

Panel A The number of SIC peers and RPE

	Dependent variable: ln(Δ*Total Compensation*)				
Independent Variables	(1)	(2)	(3)	(4)	(5)
		# of SIC Peers			
	Full Sample	Low	Med	High	High - Low
Firm Return	0.324***	0.351***	0.357***	0.320***	-0.031
	(14.956)	(9.131)	(9.100)	(9.771)	(0.539)
***Peer Return* (*SIC*)**	**-0.113*** **	**-0.058**	**-0.080**	**-0.190*** **	**-0.131**
	(-3.357)	**(-1.353)**	**(-1.367)**	**(-2.873)**	**(0.110)**
Test strong RPE F-Stat	**29.94**	**26.20**	**16.27**	**3.88**	
($\beta_1 + \beta_2 = 0$)	**(0.00)**	**(0.00)**	**(0.00)**	**(0.05)**	
Control variables	Included	Included	Included	Included	
Industry-Year FE	Yes	Yes	Yes	Yes	
# of observations	18,698	7,417	5,071	6,210	
Adjusted R-squared	0.064	0.091	0.106	0.072	

Panel B The number of GICS peers and RPE

		# of GICS Peers			
	Full Sample	Low	Med	High	High - Low
Firm Return	0.320***	0.335***	0.350***	0.304***	-0.031
	(15.540)	(9.447)	(9.246)	(8.759)	(0.539)
***Peer Return* (*GICS*)**	**-0.104*** **	**-0.014**	**-0.174*** **	**-0.155** **	**-0.141* **
	(-3.101)	**(-0.313)**	**(-2.580)**	**(-2.322)**	**(0.086)**
Test strong RPE F-Stat	**34.38**	**40.82**	**5.70**	**4.83**	
($\beta_1 + \beta_2 = 0$)	**(0.00)**	**(0.00)**	**(0.02)**	**(0.03)**	
Control variables	Included	Included	Included	Included	
Industry-Year FE	Yes	Yes	Yes	Yes	
# of observations	18,620	6,600	6,098	5,757	
Adjusted R-squared	0.075	0.090	0.085	0.075	

Table 5 Comparison with alternative industry classifications

This table presents the results obtained from the regression of the natural logarithm of the change in total CEO flow compensation on own firm stock returns (*Firm Return*), peer returns based on TNIC (*Peer Return*), either peer returns based on SIC (*Peer Return* (SIC)) in column (1) and (3) or peer returns based on GICS (*Peer Return* (GICS)) in column (2) and (4), control variables, industry-year fixed effects, and year fixed effect during sample period 1996-2011. In column (1) and (2), we do not include industry fixed effects. In column (3) and (4), we include industry fixed effects for industries in which the regression uses. All variables are winsorized at 1% and 99% levels and are defined in the Appendix A. Standard errors are clustered by firm. ***, **, and * represent significance level at the 1%, 5%, and 10% level, respectively. Robust t-statistics are in parentheses.

	Dependent variable: ln(Δ*Total Compensation*)			
Independent variables	(1)	(2)	(3)	(4)
Firm Return	0.324***	0.325***	0.327***	0.326***
	(16.814)	(16.858)	(16.632)	(16.683)
Peer Return	**-0.067****	**-0.082*****	**-0.069****	**-0.085*****
	(-2.557)	**(-3.052)**	**(-2.544)**	**(-3.080)**
Peer Return (*SIC*)	-0.027	-	-0.029	-
	(-1.125)	-	(-1.162)	-
Peer Return (*GICS*)	-	-0.006	-	-0.006
	-	(-0.223)	-	(-0.205)
Δ*Size*	0.307***	0.320***	0.314***	0.330***
	(10.779)	(11.102)	(10.466)	(10.802)
Δ*Book-to-Market*	0.071***	0.079***	0.076***	0.079***
	(3.117)	(3.481)	(3.253)	(3.443)
Idiosyncratic Volatility	-0.158**	-0.138*	0.073	0.043
	(-2.015)	(-1.783)	(0.772)	(0.464)
ln(*CEO Tenure*)	0.014***	0.014***	0.015***	0.013**
	(2.862)	(2.700)	(2.907)	(2.532)
Ownership	-0.104	-0.096	-0.075	-0.081
	(-1.530)	(-1.425)	(-0.994)	(-1.139)
Constant	0.017	0.014	0.009	0.058
	(1.404)	(1.213)	(0.168)	(1.073)
Industry Fixed Effects	No	No	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
# of observations	18,698	18,620	18,698	18,620
Adj. R-squared	0.059	0.060	0.049	0.054

Table 6 Dynamic Peer Groups and RPE tests

This table presents the results obtained from the regression of the natural logarithm of the change in total CEO flow compensation on own firm stock returns (*Firm Return*), peer returns based on TNIC (*Peer Return*), control variables, and industry-year fixed effect during sample period 1996-2011. In Panel A, the top row denote the way how we define peer firms. Past peers in column (1) are firms that were in the same TNIC group as the focal firm in period *t-1* but not in period *t* any more. Expected exit peers in column (2) are firms that are in the same TNIC group as the focal firm in period *t* but will not be in period *t+1*. Current peers in column (3) are defined as firms that are in the same TNIC group as the focal firm in period *t-1* as well as in period *t*. New peers in column (4) are firms that were not in the same TNIC group in period *t-1* but are in the same TNIC group in period *t*. Lastly, we define future peers in column (5) and (6) as firms that are not in the same TNIC group as the focal firm in period *t* but will be in the same TNIC group in period *t+1* or *t+2*. We then calculate equal-weighted average stock returns of each set of peers in period t and replace these the *Peer Return* variable in equation (1) with each of these calculated stock returns. In Panel B, we compare the estimated coefficients by simultaneously including peer return variables that are constructed based on different definition of peer firms. All variables are winsorized at 1% and 99% levels and are defined in the Appendix. Standard errors are clustered by firm. ***, **, and * represent significance level at the 1%, 5%, and 10% level, respectively. Robust t-statistics are in parentheses.

Panel A: Peer Dynamics

	Dependent variable: ln(Δ*Total Compensation*)					
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
	Past	Exp. Exit	Current	New	Future t+1	Future t+2
Firm Return	0.317***	0.310***	0.331***	0.318***	0.304***	0.292***
	(13.340)	(13.217)	(15.174)	(14.295)	(12.269)	(10.584)
Peer Return	**-0.040**	**-0.088*** **	**-0.129*** **	**-0.100*** **	**-0.085*** **	**-0.054**
	(-1.355)	**(-2.638)**	**(-3.426)**	**(-3.047)**	**(-2.584)**	**(-1.581)**
Δ*Size*	0.266***	0.282***	0.275***	0.278***	0.284***	0.274***
	(7.202)	(7.540)	(7.916)	(7.708)	(7.384)	(6.242)
Δ*Book-to-Market*	0.094***	0.075**	0.087***	0.082***	0.063**	0.078**
	(3.248)	(2.516)	(3.329)	(3.038)	(2.042)	(2.256)
Idiosyncratic Volatility	-0.040	-0.148	-0.008	-0.081	-0.160	-0.309**
	(-0.335)	(-1.211)	(-0.082)	(-0.741)	(-1.289)	(-1.963)
ln(*CEO Tenure*)	0.014**	0.012*	0.013**	0.015**	0.015**	0.017**
	(2.209)	(1.867)	(2.303)	(2.464)	(2.232)	(2.240)
Ownership	0.034	0.039	-0.005	-0.033	-0.002	0.018
	(0.343)	(0.434)	(-0.054)	(-0.364)	(-0.021)	(0.165)
Constant	-0.003	0.010	0.005	0.004	0.012	0.009
	(-0.158)	(0.563)	(0.306)	(0.251)	(0.708)	(0.447)
Industry-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
# of observations	15,048	14,733	17,262	15,987	14,069	11,567
Adjusted R-squared	0.077	0.080	0.080	0.082	0.080	0.070

Panel B: Coefficient comparison

	Dependent variable: ln(Δ*Total Compensation*)				
Independent variables	(1)	(2)	(3)	(4)	(5)
Firm Return	0.327***	0.320***	0.316***	0.308***	0.300***
	(14.390)	(13.272)	(13.234)	(12.091)	(10.535)
***Peer Return* (*Current*)**	**-0.075***	**-0.086***	-	-	-
	(-1.723)	**(-1.704)**	-	-	-
***Peer Return* (*New*)**	**-0.061***	**-**	**-0.107****	**-0.085****	**-0.076***
	(-1.851)	**-**	**(-2.480)**	**(-2.203)**	**(-1.768)**
Peer Return (*Expected to Exit*)	-	-0.038	-0.012	-	-
	-	(-1.039)	(-0.295)	-	-
Peer Return (*Future t+1*)	-	-	-	-0.049	-
	-	-	-	(-1.403)	-
Peer Return (*Future t+2*)	-	-	-	-	-0.020
	-	-	-	-	(-0.542)
Δ*Size*	0.274***	0.279***	0.275***	0.282***	0.263***
	(7.598)	(7.454)	(7.178)	(7.134)	(5.897)
Δ*Book-to-Market*	0.081***	0.075**	0.078***	0.060*	0.073**
	(3.015)	(2.536)	(2.587)	(1.922)	(2.045)
Idiosyncratic Volatility	-0.079	-0.141	-0.170	-0.208	-0.333**
	(-0.721)	(-1.159)	(-1.354)	(-1.608)	(-2.073)
ln(*CEO Tenure*)	0.015**	0.012*	0.014**	0.016**	0.017**
	(2.430)	(1.844)	(2.115)	(2.320)	(2.184)
Ownership	-0.029	0.044	-0.005	-0.017	0.008
	(-0.318)	(0.483)	(-0.050)	(-0.177)	(0.069)
Constant	0.007	0.014	0.012	0.017	0.012
	(0.437)	(0.782)	(0.669)	(0.908)	(0.569)
Industry-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
# of observations	15,987	14,733	14,182	13,362	11,053
Adjusted R-squared	0.083	0.080	0.079	0.080	0.071

Table 7 Cross-sectional tests: Strategic Flexibility and RPE

This table presents the results obtained from the regression of the natural logarithm of the change in total CEO flow compensation on own firm stock returns (*Firm Return*), peer returns based on TNIC (*Peer Return*), control variables, and industry-year fixed effect during sample period 1996-2011. All variables are winsorized at 1% and 99% levels and are defined in the Appendix A. In column (1) and (2), firm-year observations are divided into two groups based on the median value of market-to-book ratio. In column (3) and (4), observations are divided into two groups based on the median value of peer-adjusted stock return in period *t-1*. In column (5) and (6), observations are divided into two groups if the firm's asset growth in period *t+1* is greater than zero (i.e., High group), zero otherwise (i.e., Low group). Standard errors are clustered by firm. ***, **, and * represent significance level at the 1%, 5%, and 10% level, respectively. Robust t-statistics are in parentheses.

	Dependent variable: ln(Δ*Total Compensation*)					
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
	Market-to-Book		*Peer Adjusted Return*		*Asset Growth*	
	Low	High	Low	High	Low	High
Firm Return	0.336***	0.336***	0.323***	0.355***	0.366***	0.292***
	(11.423)	(10.654)	(10.547)	(10.318)	(12.295)	(9.834)
Peer Return	**-0.180*** **	**-0.064**	**-0.218*** **	**-0.111*** **	**-0.192*** **	**-0.075** **
	(-4.619)	**(-1.391)**	**(-4.705)**	**(-2.587)**	**(-4.013)**	**(-1.999)**
Δ*Size*	0.196***	0.361***	0.186***	0.256***	0.187***	0.324***
	(4.472)	(6.693)	(3.425)	(5.215)	(3.983)	(6.258)
Δ*Book-to-Market*	0.058**	0.096	0.059*	0.096**	0.095***	0.039
	(1.962)	(1.630)	(1.747)	(2.146)	(2.687)	(1.134)
Idiosyncratic Volatility	0.008	-0.097	0.172	-0.196	-0.191	0.299*
	(0.048)	(-0.566)	(0.933)	(-0.960)	(-1.143)	(1.701)
ln(*CEO Tenure*)	0.007	0.022**	0.029***	0.001	0.024***	0.012
	(0.921)	(2.393)	(2.786)	(0.139)	(2.635)	(1.204)
Ownership	-0.002	-0.066	-0.094	0.028	-0.192	0.029
	(-0.012)	(-0.557)	(-0.651)	(0.174)	(-1.515)	(0.198)
Constant	0.036*	-0.029	-0.048*	0.055**	0.011	-0.018
	(1.653)	(-1.198)	(-1.815)	(1.997)	(0.462)	(-0.675)
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
# of observations	9,551	9,551	8,514	8,514	10,037	9,068
Adj. R-squared	0.105	0.085	0.086	0.096	0.083	0.112
Coefficient difference (High - Low)						
Δ Coeff.	0.116**		0.107*		0.117**	
P-value	(0.049)		(0.072)		(0.047)	